Integrative Cancer Research Workspace
Genome Annotation
Special Interest Group
Mission Statement
DRAFT

Opportunities / Needs we will address:
The mission of the Genome Annotation SIG is to provide data and tools that will greatly enhance the cancer research community's access to high quality, comprehensive gene annotations. Having standardized access to these data sources will support a consistent view of all available gene information. This will be achieved by adapting existing software that meets the needs of the user community to comply with caBIG and by creating new software and tools.

The interests of this SIG are wide-ranging and include:
- Generation of automated, computed annotation
- Mapping annotations onto ontologies
- Mapping objects between systems (matching indexes) & linking associated objects (genes, proteins, pathways, SNPs, etc) to each other
- Integration applications that combine caBIG experimental data with annotations
- Collection and exchange of externally curated annotations
- Collection and exchange of "active" annotation or feedback from user community

The Genome Annotation SIG software will be able to associate high throughput experimental data being stored and shared within caBIG with both computed and manually curated annotation of the genes/proteins being assayed.

The annotation should be inter-operable between systems to allow for comparisons and exchange of annotations between applications.

What are we doing/will do to address these needs:
The Genome Annotation SIG will create collaborations between developers and adopters so that development addresses true needs of the user community.

In the pilot phase, each of the current developer centers will be adapting their annotation tools to use caBIG API, objects and standards to facilitate annotation exchange and incorporation into other caBIG compatible tools.

As a group, this SIG will address critical issues such as choice of identifiers, naming conventions & synonyms, overlap and redundancy of annotations from alternative sources, propagation of annotation errors, mapping between genomes, genes, transcripts and proteins, reliability scoring for annotations, etc.

We will keep abreast of annotation efforts in the field, such as the ENCODE project, to ensure that key data sources and conventions are represented and to avoid duplication of efforts.

Principles guiding our work:
The functionality covered by end-user software produced by this SIG will be based on needs of the cancer research community in general. caBIG adopters will reflect and communicate this need to steer development towards unmet needs.